

On the univariate representation of multivariate volatility models with common factors

Citation for published version (APA):

Hecq, A. W., Laurent, S. F. J. A., & Palm, F. C. (2011). *On the univariate representation of multivariate volatility models with common factors*. METEOR, Maastricht University School of Business and Economics. METEOR Research Memorandum No. 011 <https://doi.org/10.26481/umamet.2011011>

Document status and date:

Published: 01/01/2011

DOI:

[10.26481/umamet.2011011](https://doi.org/10.26481/umamet.2011011)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Alain Hecq, Sébastien Laurent, Franz Palm

**On the Univariate Representation
of Multivariate Volatility Models
with Common Factors**

RM/11/011



METEOR

Maastricht University School of Business and Economics
Maastricht Research School of Economics
of Technology and Organization

P.O. Box 616
NL - 6200 MD Maastricht
The Netherlands

ON THE UNIVARIATE REPRESENTATION OF MULTIVARIATE VOLATILITY MODELS WITH COMMON FACTORS

Alain Hecq, Sébastien Laurent and Franz C. Palm*

Department of Quantitative Economics

Maastricht University

The Netherlands

November 12, 2010

Abstract

First, we investigate the minimal univariate representation of some well known n -dimensional conditional volatility models. Simple systems (e.g. a $VEC(0,1)$) for the joint behaviour of several variables imply individual processes with a lot of persistence in the form of long order lags. We show that in the presence of factors, parsimonious univariate representations (e.g. $GARCH(1,1)$) can result from large multivariate models generating the conditional variances and conditional correlations.

Second, we propose an approach to use empirical results for these univariate processes in the analysis of the underlying multivariate, possibly high-dimensional, GARCH process. We use reduced rank procedures to discriminate between a system with seemingly unrelated assets (e.g. a diagonal model) from a set of series with few common sources of volatility. Among the analyzed procedures, the canonical correlation test statistics on logs of squared returns proposed by Engle and Marcucci (2006) has quite good properties even in the case of falsely omitted cross-moments. Out of 30 returns from the NYSE, six returns are shown to display a parsimonious $GARCH(1,1)$ model for their conditional variance. We do not reject the hypothesis that a single common volatility factor drives these six series.

JEL: C10, C32

Keywords: Co-movements, factor models, ARMA, GARCH, VEC, Final equations, Long memory.

*Corresponding author: Franz Palm, Department of Quantitative Economics, Maastricht University, P.O.Box 616, 6200 MD Maastricht, The Netherlands. E-mail: f.palm@maastrichtuniversity.nl.

1 Motivation

Most financial econometrics textbooks start with univariate models to explain the presence of a time-varying conditional volatility pattern in asset returns. Then the properties of different models (e.g. GARCH, EGARCH, realized volatility, stochastic volatility,...) are discussed together with their most important features and their relation with stylized facts observed in financial data. It is also explained that to correctly account for the link between several series (e.g. contagion effects) and to study time-varying conditional correlations, a multivariate framework should be used. Indeed, understanding and predicting the dependence in the second order moments of asset returns is important for many issues in financial econometrics and management.

Unrestricted multivariate GARCH models often suffer from the curse of dimensionality. An n -dimensional unrestricted VEC(0, q) already implies $((n^2 + n)/2)^2 q$ unknown coefficients plus n intercepts. As for reason of diversification most portfolios involve a large number of assets, the need for more parsimonious forms is obvious. Let us just name the diagonal model, the constant conditional correlation (CCC), the dynamic conditional correlation (DCC), the dynamic equicorrelation (DECO, see Engle and Kelly, 2008), the approach by Baba, Engle, Kraft and Kroner (1999, hereafter BEKK), the orthogonal GARCH or factor GARCH models as examples proposed in the literature to restrict the multivariate setting towards a manageable size as well as to impose the positive definitiveness of the covariance matrix (see *inter alia* the surveys by Bauwens, Laurent and Rombouts (2006) or Silvennoinen and Terasvirta (2009)). To make their estimation feasible on large portfolios some of these models impose very strong restrictions on the dynamics of the covariance or correlation. One simple way to relax these constraints is to assume a block-triangular structure as Engle and Marcucci (2006) implicitly do to obtain the pure variance model or to assume a block-diagonal structure where the dynamics is constrained to be equal only within groups of variables (see for instance the BLOCK-DCC and BLOCK-DECO of respectively Billio, Caporin and Gobbo, 2003 and Engle and Kelly, 2008).

Another route is taken in this paper. Indeed we revert to some extent the problem and we focus our analysis on a group of assets that must be jointly analyzed. To do so, we compute the final equation representation of multivariate GARCH models using a framework similar to that in Zellner and Palm (see *inter alia* 1974, 1975, 2004) for the conditional mean. This framework allows us to derive the marginal GARCH representation for the conditional variances and conditional covariances of the multivariate GARCH model (e.g. Nijman and Sentana, 1996). However this implied univariate representation is often far from being parsimonious. There consequently exists a paradox between theoretical marginal volatility models derived from a multivariate model and empirical findings, a paradoxical issue similar to the one observed for the VAR(p) and the marginal implied ARMA models (see Cubadda, Hecq and Palm, 2008, 2009).

Indeed in empirical work, estimated univariate GARCH models are, too a large extent, very parsimonious. The GARCH(1,1) specification for instance is able to capture the time varying volatility in second moments of many symmetric asset returns. Consequently we look at some multivariate representations that could at least be compatible with the univariate structure obtained in empirical work (like a GARCH(1,1)). We show that a familiar multivariate model such as the “unrestricted” BEKK, generally does not imply low order dynamics for the volatility of individual asset returns. Therefore, we extend the analysis that has been developed by Cubadda

et al. (2009) for vector autoregressive models to factor representations in the second moments. These are the factor GARCH specification of Engle et al. (1990) and the Engle and Marcucci (2006) factor pure variance model. These two specifications allow to obtain such parsimonious univariate representations for potentially large multivariate systems. This also implies that obtaining a very parsimonious univariate representation of individual returns might be an indication of co-movements in the volatility while on the opposite, finding less parsimonious univariate schemes is interpreted as a sign of absence of co-movements.

These common volatility systems are not the only ones yielding parsimonious GARCH orders. As illustrated in Section 2, a low order diagonal model which implies no contagion effects leads to similar low order univariate GARCH models. This is also the case for the DCC which imposes a block-diagonal structure on its VEC representation. A related paper to ours is Nijman and Sentana (1996). Indeed these authors also study the marginal models derived from a multivariate GARCH process with a particular focus on the aggregation of individual series. They also sketch that the presence of a factor model in volatility might be important in the marginalization of multivariate systems but without giving the orders of the univariate representations nor the proofs. These are the points we look at more closely in Sections 2 and 3.

The objective of this paper is twofold. First, we propose in Section 3 some multivariate models accounting for co-movements and show that the implied marginal volatility processes are of low order. In particular we are able to discriminate between a diagonal model and a general multivariate framework, with correlated conditional variances and contagion effects, driven by a small number of common factors in volatility. Second, we present in Section 4 tests for the detection of co-movements in volatility and study their small sample properties. It is shown that the approach proposed by Engle and Marcucci (2006) for the logarithms of squared returns performs quite well although it omits the cross-moments. In Section 5 we apply our findings to a set of thirty financial assets. We determine their univariate volatility processes and group these processes into parsimonious and non parsimonious ones. Then we find evidence for the presence of co-movements between the six series with low order univariate volatility schemes. The final section concludes.

2 The final equation representation of multivariate GARCH models

To state the notation for univariate processes, ε_{1t} is such that $\varepsilon_{1t} = u_{1t}\sqrt{h_{11t}}$, where u_{1t} has any centered parametric distribution with unitary variance and the conditional variance follows a GARCH(p, q) with $h_{11t} = \omega_1 + \sum_{j=1}^p \beta_{1,j} h_{11t-j} + \sum_{i=1}^q \alpha_{1,i} \varepsilon_{1t-i}^2$. Consequently here, p refers to the GARCH terms and q is the order of the moving average ARCH term. The error in the squared returns is obtained as usual using $v_{1t} = \varepsilon_{1t}^2 - h_{11t}$. In a GARCH(1, 2) that we take on purpose to emphasize different orders p and q , substituting for the variances in the variance equation and rearranging terms we can write our model in terms of the squared errors: $\varepsilon_{1t}^2 = \omega_1 + (\alpha_{1,1} + \beta_{1,1})\varepsilon_{1t-1}^2 + \alpha_{1,2}\varepsilon_{1t-2}^2 + v_{1t} - \beta_{1,1}v_{1t-1}$ (see *inter alia* Bollerslev, 1996). Thus, the squared errors follow a heteroskedastic weak GARCH univariate ARMA(2, 1) process. In general a GARCH(p, q) has got an ARMA(max(p, q), p) representation for the squared errors. In most of the examples we have considered, we take $q \geq p$.

For multivariate modeling, we denote $\varepsilon_t = H_t^{1/2} z_t$, $t = 1, \dots, T$, the $(n \times 1)$ vector of excess returns¹ of financial assets observed at the time period t . T is the number of observations and z_t is assumed to be an i.i.d. random variable with mean 0 and variance I_n such that H_t is the conditional covariance matrix of the n assets.

To simplify the presentation, let us use a bivariate GARCH in its extensive VEC(0, 1) representation

$$\begin{aligned} h_{11t} &= \omega_1 + \alpha_{11}\varepsilon_{1t-1}^2 + \alpha_{12}\varepsilon_{2t-1}^2 + \alpha_{13}\varepsilon_{1t-1}\varepsilon_{2t-1} \\ h_{12t} &= \omega_2 + \alpha_{21}\varepsilon_{1t-1}^2 + \alpha_{22}\varepsilon_{2t-1}^2 + \alpha_{23}\varepsilon_{1t-1}\varepsilon_{2t-1} \\ h_{22t} &= \omega_3 + \alpha_{31}\varepsilon_{1t-1}^2 + \alpha_{32}\varepsilon_{2t-1}^2 + \alpha_{33}\varepsilon_{1t-1}\varepsilon_{2t-1}, \end{aligned} \quad (1)$$

where usual non-negativity and stationarity restrictions on the parameters are assumed to be met. h_{iit} is the conditional variance of asset $i = 1, 2$ and h_{ijt} stands for the conditional covariance for $i \neq j$. Let us write the system (1) in terms of observed squared returns and covariances

$$\begin{pmatrix} \varepsilon_{1t}^2 \\ \varepsilon_{1t}\varepsilon_{2t} \\ \varepsilon_{2t}^2 \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} \varepsilon_{1t-1}^2 \\ \varepsilon_{1t-1}\varepsilon_{2t-1} \\ \varepsilon_{2t-1}^2 \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}, \quad (2)$$

with $v_{1t} = \varepsilon_{1t}^2 - h_{11t}$, $v_{2t} = \varepsilon_{1t}\varepsilon_{2t} - h_{12t}$, $v_{3t} = \varepsilon_{2t}^2 - h_{22t}$, or simply for $\epsilon_t = \{\varepsilon_{1t}^2, \varepsilon_{1t}\varepsilon_{2t}, \varepsilon_{2t}^2\}'$

$$\alpha(L)\epsilon_t = \omega + v_t, \quad (3)$$

with the lag operator L such that $Lz_t = z_{t-1}$. $\alpha(L) = (I - \alpha L)$ is a matrix polynomial of degree one in this example, where α is the matrix of coefficients. Note that (3) has the Wold representation $\epsilon_t = \alpha^{-1}(L)\omega + \alpha^{-1}(L)v_t$ of squared disturbances and cross products with v_t being a martingale difference stationary process as $E(v_t|\Omega_{t-1}) = 0$, $E(v_tv'_{t-i}|\Omega_{t-i}) = 0$, $i > 0$ and Ω_{t-i} being the past of v_t up to and including period $t - i$. The vector v_t is therefore serially uncorrelated. The dimension of vectors ϵ_t , ω and v_t is $N = (n^2 + n)/2$ and consequently there are N^2q unknown parameters in $\alpha(L)$ for a VEC(0, q).

Let us now premultiply both sides of (3) by the adjoint of the matrix polynomial $\alpha(L)$ to obtain

$$Adj\{\alpha(L)\}\alpha(L)\epsilon_t = Adj\{\alpha(L)\}\omega + Adj\{\alpha(L)\}v_t, \quad (4)$$

or²

$$\begin{aligned} |\alpha(L)|\epsilon_t &= \omega^* + Adj\{\alpha(L)\}v_t \\ &= \omega^* + \theta(L)\zeta_t, \end{aligned} \quad (5)$$

¹By excess return we mean that the conditional mean (a constant or an ARMA model for instance) has been subtracted from the returns r_t , say $\varepsilon_t = r_t - E(r_t|\Omega_{t-1})$, where Ω_{t-1} denotes the information set up to and including $t - 1$. So generally, it is a martingale difference sequence, not an i.i.d. random variable.

²Alternatively this can be shown from $\alpha(L)\epsilon_t = \omega + v_t$ that leads to $\epsilon_t = \alpha(L)^{-1}\omega + \alpha(L)^{-1}v_t$. Using the definition $\alpha(L)^{-1} = Adj\{\alpha(L)\}/\det[\alpha(L)]$ we end up with the same result. However strictly speaking this necessitates conditions on the invertibility of $\alpha(L)$.

where $|\alpha(L)| = \det\{\alpha(L)\}$, i.e. the determinant of the matrix polynomial $\alpha(L)$, is a scalar polynomial in L ; $Adj\{\alpha(L)\}$ denotes the adjoint (or the adjugate) of $\alpha(L)$ and $\omega^* = Adj\{\alpha(L)\}\omega$. In a $VEC(0, q)$, each component of the vector ϵ_t has a weak $ARMA(Nq, (N-1)q)$ or Wold representation that can be written as follows for instance for the first element ε_{1t}^2

$$\begin{aligned} |\alpha(L)|\varepsilon_{1t}^2 &= \omega_1^* + \theta_1'(L)\zeta_t, \\ &= \omega_1^* + \beta_1(L)u_{1t}, \end{aligned} \tag{6}$$

where $\theta_1'(L)$ is the first row of $\theta(L)$ and $u_{1t} = \varepsilon_{1t}^2 - h_{11t}$ is a white noise with h_{11t} being the linear projection of ε_{1t}^2 on the space generated by the past of ε_{1t}^2 . The univariate weak GARCH representation of h_{11t} is obtained by substituting $\varepsilon_{1t}^2 - h_{11t}$ for u_{1t} .

Proposition 1 (see also Nijman and Sentana, 1996) summarizes the main features of the final equation representation (5).

Proposition 1 *In a n -dimensional $VEC(0, q)$, each univariate component is weakly GARCH with a univariate $ARMA(Nq, (N-1)q)$ representation of the squared returns and cross returns with the same value of autoregressive parameters, $N = (n^2 + n)/2$. Consequently each component follows a weak $GARCH((N-1)q, Nq)$. The orders should be taken as upperbounds for the orders of the univariate ARMA and GARCH models.*

Proof. The proof is obvious from the definition of the determinant and the adjoint. This well known result is simply due to the fact that in the $VEC(0, q)$ for instance $|\alpha(L)|$ contains by construction up to L^{Nq} terms and the adjoint matrix is a collection of $\{(N-1) \times (N-1)\}$ cofactor matrices, each of the matrix elements can contain the terms $1, L, \dots, L^q$. As v_t is a vector martingale difference sequence it is serially uncorrelated and each element of $Adj\{\alpha(L)\}v_t$ can be represented as a univariate moving average and therefore it is a weak GARCH process. ■

Proposition 2 *Proposition 1 generalizes in a straightforward manner to show that for a n -dimensional $VEC(p, q)$, each univariate component (squared and cross returns) has a univariate $ARMA(N \max\{p, q\}, (N-1) \max\{p, q\} + p)$ representation at most with the same value of autoregressive parameters. Consequently each component follows a weak $GARCH((N-1) \max\{p, q\} + p, N \max\{p, q\})$ process at most with $N = (n^2 + n)/2$.*

Proof. It is similar to that of Proposition 1. ■

The above outcomes, that apply the usual results of the $VAR(p)$ and $VARMA(p, q)$ are generally not in agreement with empirical findings suggesting low order univariate GARCH schemes. Indeed, for $n = 20$ assets, a $VEC(0, 2)$ implies individual $ARMA(420, 418)$ models in squared returns and cross-products and individual $GARCH(418, 420)$ processes. Obviously these orders should be taken as upperbounds. For instance in the example with the $VEC(0, 1)$ in (2) for $n = 2$, the determinant of $\alpha(L)$ is

$$\begin{aligned} |\alpha(L)| &= (\alpha_{11}\alpha_{23}\alpha_{32} + \alpha_{12}\alpha_{21}\alpha_{33} - \alpha_{12}\alpha_{31}\alpha_{23} - \alpha_{21}\alpha_{13}\alpha_{32} + \alpha_{13}\alpha_{22}\alpha_{31} - \alpha_{11}\alpha_{22}\alpha_{33})L^3 \\ &\quad + (\alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21} + \alpha_{11}\alpha_{33} - \alpha_{13}\alpha_{31} + \alpha_{22}\alpha_{33} - \alpha_{23}\alpha_{32})L^2 \\ &\quad - (\alpha_{11} + \alpha_{22} + \alpha_{33})L + 1, \end{aligned}$$

and the adjoint minus the identity (to save space), i.e. $Adj\{\alpha(L)\} - I_N$, is

$$\begin{pmatrix} (\alpha_{22}\alpha_{33}-\alpha_{23}\alpha_{32})L^2-(\alpha_{33}+\alpha_{22})L & (\alpha_{13}\alpha_{32}-\alpha_{12}\alpha_{33})L^2+\alpha_{12}L & (\alpha_{12}\alpha_{23}-\alpha_{13}\alpha_{22})L^2+\alpha_{13}L \\ (\alpha_{31}\alpha_{23}-\alpha_{21}\alpha_{33})L^2+\alpha_{21}L & (\alpha_{11}\alpha_{33}-\alpha_{13}\alpha_{31})L^2-(\alpha_{33}+\alpha_{11})L & (\alpha_{21}\alpha_{13}-\alpha_{11}\alpha_{23})L^2+\alpha_{23}L \\ (\alpha_{21}\alpha_{32}-\alpha_{22}\alpha_{31})L^2+\alpha_{31}L & (\alpha_{12}\alpha_{31}-\alpha_{11}\alpha_{32})L^2+\alpha_{32}L & (\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})L^2-(\alpha_{22}+\alpha_{11})L \end{pmatrix}.$$

Therefore, each squared excess return and cross-products follows an ARMA(3,2) model at most, implying thus GARCH(2,3) models for conditional variances and the conditional covariances. However, each element in the power of L being the sum of the products of coefficients smaller than one and possibly zero (e.g. $\alpha_{11}\alpha_{22}\alpha_{33}$), a lower order structure can be identified in small samples. There might also exist coincidental situations (Granger and Newbold, 1986) in which there exist “quasi” common roots in the determinant and the adjoint (see Nijman and Sentana, 1996 for an example). A particular case in which there are exact common roots between the implied AR and MA parts is the diagonal model of Bollerslev (1990) where

$$(I - \alpha L) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \alpha_{11} & 0 & 0 \\ 0 & \alpha_{22} & 0 \\ 0 & 0 & \alpha_{33} \end{pmatrix} L. \quad (7)$$

Hence a diagonal multivariate strong GARCH process is identical to a set of strong GARCH univariate processes with possibly contemporaneous correlated disturbances.³ Remark that we only work on marginalization of processes and we do not extend these results to the aggregation as in Nijman and Sentana (1996).

Another popular multivariate GARCH model is the BEKK(p, q) of Baba et al. (1989). In this specification there are no common roots between the determinant and the adjoint and consequently the general results of Proposition 1 apply. Indeed, let us take a BEKK(0, 1) representation for two time series

$$H_t = \Gamma'_0 \Gamma_0 + \Gamma'_1 \varepsilon_{t-1} \varepsilon'_{t-1} \Gamma_1,$$

where $\omega_i = [\Gamma'_0 \Gamma_0]_{ii}$ and with the elements of Γ_1 denoted γ_{ij} . In terms of observed squared returns and covariances using the same manipulation we used before, we obtain

$$(I - \alpha L) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - \begin{pmatrix} \gamma_{11}^2 & 2\gamma_{11}\gamma_{21} & \gamma_{21}^2 \\ \gamma_{11}\gamma_{12} & (\gamma_{21}\gamma_{12} + \gamma_{11}\gamma_{22}) & \gamma_{21}\gamma_{22} \\ \gamma_{12}^2 & 2\gamma_{12}\gamma_{22} & \gamma_{22}^2 \end{pmatrix} L. \quad (8)$$

Computing the final equations representations under the restrictions (8), it emerges that the determinant and the adjoint of the matrix polynomial (8) are respectively of lag orders 3 and 2. Consequently this BEKK(0, 1), like the unrestricted VEC(0,1), generates GARCH(2, 3) univariate processes for the conditional variances and covariances. No further reduction occurs. General results that have been derived for the multivariate GARCH(p, q) apply here.

³The conditional orthogonal model further assumes that the cross-product term $\alpha_{22} = 0$ (and then $\omega_2 = 0$).

3 Factors GARCH models

Simple models (e.g. BEKK) do not imply parsimonious low order univariate GARCH processes. Independence and non-contagion of the volatility, a feature that seems very unlikely, are able to explain it. Let us now look at various factor models. Indeed, Cubadda, Hecq and Palm (2008, 2009) have derived the cases under which a co-movement or factor structure in a VAR is able to explain the parsimony of marginal ARMA models. This section gives the implications of the presence of a factor structure in second moments for the volatility observed in financial assets. In summary and anticipating the results, it will be shown that the presence of co-movements in the volatility might explain the gap between the theoretical orders and the empirical findings.

Next, we put forward a strategy consisting in using information on individual series and then testing for the presence of co-movements in order to determine the set of assets that must be jointly modeled.

3.1 The factor GARCH

Let us denote a BEKK(p, q) model with k factors F-BEKK(p, q, k) and in particular a F-BEKK(0, 1, 1) such that

$$\begin{aligned}\varepsilon_t &= H_t^{1/2} z_t \\ H_t &= \Gamma'_0 \Gamma_0 + \gamma \varphi' \varepsilon_{t-1} \varepsilon_{t-1}' \gamma',\end{aligned}$$

where in a bivariate system $\varphi' = (\varphi_1 : \varphi_2)$, $\gamma' = (\gamma_1 : \gamma_2)$ and hence $\text{rank}(\varphi \gamma') = 1$ and where $\gamma'_\perp \Gamma'_0 \Gamma_0 \neq 0$. We can write this system using the half vec operator vech such that $\text{vech}(H_t) = \text{vech}(\Gamma'_0 \Gamma_0) + \text{Avech}(\varepsilon_{t-1} \varepsilon_{t-1}')$ or in terms of squared errors and cross products $\text{vech}(\varepsilon_t \varepsilon_t') = \text{vech}(\Gamma'_0 \Gamma_0) + \text{Avech}(\varepsilon_{t-1} \varepsilon_{t-1}') + v_t$ with the multivariate martingale difference sequence $v_t = \text{vech}(\varepsilon_t \varepsilon_t') - \text{vech}(H_t)$. Note that in the one-factor case, A can also be obtained using $\varphi \varphi' (\gamma' \varepsilon_{t-1})^2 = \gamma \varphi' \varepsilon_{t-1} \varepsilon_{t-1}' \gamma'$. Therefore, the matrix $\varphi \varphi'$ is of rank one as well as the coefficient matrix A . Consequently, A can be written in a reduced form with

$$\begin{aligned}A &= \begin{pmatrix} \gamma_1^2 \varphi_1^2 & 2\gamma_1^2 \varphi_1 \varphi_2 & \gamma_1^2 \varphi_2^2 \\ \gamma_1 \gamma_2 \varphi_1^2 & 2\gamma_1 \gamma_2 \varphi_1 \varphi_2 & \gamma_1 \gamma_2 \varphi_2^2 \\ \gamma_2^2 \varphi_1^2 & 2\gamma_2^2 \varphi_1 \varphi_2 & \gamma_2^2 \varphi_2^2 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ \frac{\gamma_2}{\gamma_1} \\ (\frac{\gamma_2}{\gamma_1})^2 \end{pmatrix} \begin{pmatrix} \gamma_1^2 \varphi_1^2 & 2\gamma_1^2 \varphi_1 \varphi_2 & \gamma_1^2 \varphi_2^2 \end{pmatrix}\end{aligned}$$

such that there exists a cofeature matrix

$$\delta = \begin{pmatrix} -\frac{\gamma_2}{\gamma_1} & -\frac{\gamma_2^2}{\gamma_1^2} \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

with $\delta' A = 0_{2 \times 3}$. For the determinant

$$\det(I - AL) = 1 - (\gamma_1 \varphi_1 + \gamma_2 \varphi_2)^2 L$$

and the adjoint

$$\begin{pmatrix} -(\gamma_2^2\varphi_2^2 + 2\gamma_1\varphi_1\gamma_2\varphi_2)L + 1 & 2L\gamma_1^2\varphi_1\varphi_2 & L\gamma_1^2\varphi_2^2 \\ L\gamma_1\gamma_2\varphi_1^2 & -(\gamma_1^2\varphi_1^2 + \gamma_2^2\varphi_2^2)L + 1 & L\gamma_1\gamma_2\varphi_2^2 \\ L\gamma_2^2\varphi_1^2 & 2L\gamma_2^2\varphi_1\varphi_2 & -(\gamma_1^2\varphi_1^2 + 2\gamma_2\varphi_2\gamma_1\varphi_1)L + 1 \end{pmatrix}$$

it is seen that conditional second moments, squared returns and their cross-products are GARCH(1,1). These findings generalize as follows:

Proposition 3 *In a n -dimensional F-BEKK(0, q , k), each univariate component of the squared returns and cross-returns has a univariate ARMA(p^* , q^*) representation with the same value of autoregressive parameters. The orders of p^* and q^* are at most $(N - (N - k))q = qk$ and hence do not depend on N . As a special case, each component of a multivariate F-BEKK(0, 1, 1) follows a weak GARCH(1, 1) whatever the number of assets jointly considered.*

The propositions and the proofs are applications of the results obtained in Cubadda, Hecq and Palm (2009) for the VAR(p). We only give the proof for the *one*-factor GARCH(0, q), i.e. the F-BEKK(0, q , 1). In this case there exists an $(N \times (N - 1))$ full column rank matrix δ (with $N = \frac{n(n+1)}{2}$) such that $\delta' \text{vech}(\varepsilon_t \varepsilon_t') = \delta' v_t$ with $\delta' \varphi = 0$.

Proof. *Let us rewrite the F-BEKK(0, q , 1) as follows*

$$Q(L)x_t = e_t,$$

where $x_t = M \text{vech}(\varepsilon_t \varepsilon_t')$, $e_t = M v_t$, $Q(L) = M \Phi(L) M^{-1}$, $M' \equiv [\delta : \delta_\perp]$, δ_\perp is the orthogonal complement of δ with $\text{span}(\delta_\perp) = \text{span}(\varphi)$. Given that x_t is a non-singular linear transformation of $\text{vech}(\varepsilon_t \varepsilon_t')$, the maximum AR and MA orders of the univariate representation of the elements of $\text{vech}(\varepsilon_t \varepsilon_t')$ must be the same as those of elements of x_t . Since $M^{-1} = [\bar{\delta} : \bar{\delta}_\perp]$, where $\bar{\delta} = \delta(\delta' \delta)^{-1}$, and $\bar{\delta}_\perp = \delta_\perp(\delta_\perp' \delta_\perp)^{-1}$, we have

$$Q(L) = \begin{bmatrix} I_{(N-1)} & 0_{(N-1) \times 1} \\ \delta_\perp' \Phi(L) \bar{\delta} & \delta_\perp' \Phi(L) \bar{\delta}_\perp \end{bmatrix},$$

from which it easily follows that $\det[Q(L)] = \det[\delta_\perp' \Phi(L) \bar{\delta}_\perp]$ is a polynomial of order q . Hence, the univariate AR order of each element of $\text{vech}(\varepsilon_t \varepsilon_t')$ are at most q . To prove the order of the MA component, let $P(L)$ denotes a submatrix of $Q(L)$ that is formed by deleting one of the first $(N - 1)$ rows and one of the first $(N - 1)$ columns of $Q(L)$. We can partition $P(L)$ as follows

$$P(L) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21}(L) & P_{22}(L) \end{bmatrix}. \quad (9)$$

Now, P_{11} is a $((N - 1) \times (N - 1))$ identity matrix, P_{12} is a $((N - 1) \times 1)$ vector of zeros, $P_{21}(L)$ is a $(1 \times (N - 1))$ polynomial vector of order q , and $P_{22}(L)$ is a scalar polynomial of order q . Hence, $\det[P(L)] = \det[P_{11}] \det[P_{22}(L)]$, and therefore $\det[P(L)]$ is of order q . Since cofactors associated with the blocks of $Q(L)$ different from P_{11} are polynomials of degree not larger than q , we conclude that the univariate MA orders of each element of $\text{vech}(\varepsilon_t \varepsilon_t')$ are at most q . Using the same results, the orders of the AR and MA polynomials of the univariate processes are qk when there are k factors. ■

3.2 The pure variance model

In this model proposed by Engle and Marcucci (2006), there only exists a reduced rank structure between the squared returns such that

$$\varepsilon_t^2 = \gamma_0 + \varphi \gamma' \varepsilon_{t-1}^2 + v_t,$$

for $\varepsilon_t^2 = (\varepsilon_{1t}^2, \dots, \varepsilon_{nt}^2)'$. This is also a special case of the VEC(0,1) given in (2). In practice however, Engle and Marcucci (2006) assume an exponential form and thus take the log-transform to ensure strictly positive squared returns (see Section 4).

The orders of the individual models are easy to obtain. Indeed note that there exists an $(n \times (n - k))$ full column rank matrix δ such that $\delta' \varepsilon_t^2 = \delta' v_t$ with $\delta' \varphi = 0$. Using the same type of proof that we have proposed for the more general model it emerges that in an n -dimensional stationary GARCH(0, q) with k pure variance component, the individual ARMA processes for the squared excess returns have both AR and MA orders not larger than kq .

Engle and Marcucci (2006) explicitly ignore the cross-products. Before looking in the next sections at the results of test statistics for the pure variance model when the correct specification includes contagion effects, let us first develop the theoretical model representation. We can indeed write the N -dimensional VEC(0, q) with $N = n(n + 1)/2$ such that

$$h_t = \omega + A_q(L) \begin{pmatrix} \varepsilon_t^2 \\ x_t \end{pmatrix}$$

with x_t the vector of $\frac{n(n-1)}{2} = N - n$ cross-product elements $\varepsilon_{it}\varepsilon_{jt}$, $i \neq j$. In terms of observable series we have

$$\{I - A_q(L)\} \begin{pmatrix} \varepsilon_t^2 \\ x_t \end{pmatrix} = \omega + v_t, \quad v_t = h_t - \begin{pmatrix} \varepsilon_t^2 \\ x_t \end{pmatrix},$$

or more explicitly

$$\begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} \varepsilon_t^2 \\ x_t \end{pmatrix} = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} + \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix}, \quad (10)$$

where to simplify notations $\phi_{ij} = \phi_{ij}(L)$ are polynomial matrices of degree q in L . Similarly we denote by $|\phi_{ii}|$ and ϕ_{ij}^a respectively the determinant and the adjoint of $\phi_{ij}(L)$.

We now marginalize (10) with respect to x_t . This can be done by first computing the final equation representation of the second row of (10) using the identity $|\phi_{22}|I_{N-n} = \phi_{22}^a \phi_{22}$. Next we multiply the first row of (10) by $|\phi_{22}|$ and the new equation for x_t by ϕ_{12} . One obtains the following system of equations

$$\begin{cases} |\phi_{22}|\phi_{11}\varepsilon_t^2 + |\phi_{22}|\phi_{12}x_t = |\phi_{22}|\omega_1 + |\phi_{22}|v_{1t} \\ \phi_{12}\phi_{22}^a\phi_{21}\varepsilon_t^2 + \phi_{12}|\phi_{22}|x_t = \phi_{12}\phi_{22}^a\omega_2 + \phi_{12}\phi_{22}^a v_{2t} \end{cases}. \quad (11)$$

Remains to subtract the second equation from the first in order to eliminate x_t to get

$$\left[\underbrace{|\phi_{22}|}_{(N-n)q} \underbrace{\phi_{11}}_q - \underbrace{\phi_{12}}_q \underbrace{\phi_{22}^a}_{(N-n-1)q} \right] \varepsilon_t^2 = \omega^* + \underbrace{|\phi_{22}|}_{(N-n)q} v_{1t} - \underbrace{\phi_{12}}_q \underbrace{\phi_{22}^a}_{(N-n-1)q} v_{2t}, \quad (12)$$

with $\omega^* = |\phi_{22}|\omega_1 - \phi_{12}\phi_{22}^a\omega_2$ and where the maximal polynomial orders are reported beneath each polynomial matrices.

This implies an n -dimensional VARMA($(N-n)q+q, (N-n)q$) pure variance process for which the implied maximal univariate orders can be determined using the general rules. This gives rise to several comments:

Remark 4 *First, under the restrictions $\phi_{12}(L) = 0$, the marginal process for ε_t^2 is a pure variance model $\phi_{11}(L)\varepsilon_t^2 = \omega_1 + v_{1t}$ in the form of an autoregression as used by Engle and Marcucci (2006). This is the case of absence of Granger-causality from x_t to ε_t^2 in the sense that in the linear predictor (in the MSE sense) or projection of ε_t^2 of its own past and on that of x_t the slope coefficients of the lagged values of x_t are zero.*

Remark 5 *An important example in which the dynamics of conditional variances and covariances are assumed to be independent, i.e. $\phi_{12}(L) = 0$ and $\phi_{21}(L) = 0$, is inspired by the DCC model of Engle (2002).*

Remark 6 *If in addition to $\phi_{12}(L) = 0$, $\phi_{11}(L)$ has rank $k \leq n$, ε_t^2 is generated by a factor pure variance model.*

Remark 7 *If the rank of the $(N \times Nq)$ polynomial matrix $A_q(L)$ of the F-BEKK(0,q,k) is $k \leq n$, we have $\text{rank}[(I_n - \phi_{11}(L)) : \phi_{12}(L)] = k$ and then $\exists \delta : \delta'[(I_n - \phi_{11}(L)), \phi_{12}(L)] = 0$. Let us premultiply (12) by δ' to obtain under the null hypothesis that $k < n$*

$$\delta'|\phi_{22}|\phi_{11}\varepsilon_t^2 = \delta'|\phi_{22}|\omega_1 + \delta'|\phi_{22}|v_{1t},$$

which implies

$$\delta'|\phi_{22}|\varepsilon_t^2 = \delta'|\phi_{22}|\omega_1 + \delta'|\phi_{22}|v_{1t},$$

and after common roots cancellation

$$\delta'\varepsilon_t^2 = \delta'\omega_1 + \delta'v_{1t},$$

which means that we have the implication of the pure variance under the null of the presence of common factors.

4 Test statistic and Monte Carlo results

Let us start with the F-BEKK(0, q, k)

$$\begin{aligned}\varepsilon_t &= H_t^{1/2} z_t \\ H_t &= \Gamma_0' \Gamma_0 + \Gamma_1' \varepsilon_{t-1} \varepsilon_{t-1}' \Gamma_1 + \dots + \Gamma_q' \varepsilon_{t-q} \varepsilon_{t-q}' \Gamma_q,\end{aligned}$$

for which the objective is the determination of the number of common factors driving H_t . From $\text{vech}(H_t) = \text{vech}(\Gamma_0' \Gamma_0) + A_1 \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + A_q \text{vech}(\varepsilon_{t-q} \varepsilon_{t-q}')$ we write

$$\text{vech}(\varepsilon_t \varepsilon_t') = \text{vech}(\Gamma_0' \Gamma_0) + A_1 \text{vech}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + A_q \text{vech}(\varepsilon_{t-q} \varepsilon_{t-q}') + v_t, \quad (13)$$

or equivalently

$$\text{vech}(\varepsilon_t \varepsilon_t') = \text{vech}(\Psi) + A W_t + v_t, \quad (14)$$

where $v_t = vech(\varepsilon_t \varepsilon_t') - vech(H_t)$ is a multivariate martingale difference sequence, $A = A_1 : \dots : A_q$ is a $(n(n+1)/2 \times n(n+1)q/2)$ matrix and $W_t = (vech(\varepsilon_{t-1} \varepsilon_{t-1}') : \dots : vech(\varepsilon_{t-q} \varepsilon_{t-q}'))'$. The number of factors is now determined by $k = \text{rank}(A)$.

In the F-BEKK, neglecting the cross-products in $vech(\varepsilon_t \varepsilon_t')$ does not affect the rank of A . For this reason, one might also test the number of factors in a pure variance framework

$$\text{diag}(\varepsilon_t \varepsilon_t') = \text{diag}(\Gamma_0' \Gamma_0) + B_1 \text{diag}(\varepsilon_{t-1} \varepsilon_{t-1}') + \dots + B_q \text{diag}(\varepsilon_{t-q} \varepsilon_{t-q}') + u_t, \quad (15)$$

or in a more compact form

$$\varepsilon_t^2 = \gamma_0 + B W_t + u_t, \quad (16)$$

where $\gamma_0 = \text{diag}(\Gamma_0' \Gamma_0)$ is a $(n \times 1)$ vector of constants, $B = B_1 : \dots : B_q$ is a $(n \times nq)$ matrix, $W_t = \text{diag}(\varepsilon_{t-1} \varepsilon_{t-1}') : \dots : \text{diag}(\varepsilon_{t-q} \varepsilon_{t-q}')$, $u_t = \text{diag}(\varepsilon_t \varepsilon_t') - \text{diag}(H_t)$ and $k = \text{rank}(B)$. While the null assumptions about the rank are the same in (15) or (16), the performance of these tests might be different in small samples and consequently we consider both testing strategies in the sequel.

To test the rank of A and B , we can use the canonical correlation analysis as in Engle and Marcucci (2006), i.e.,

$$\Sigma_{YY}^{-1} \Sigma_{YW} \Sigma_{WW}^{-1} \Sigma_{WY}, \quad (17)$$

where Σ_{ij} are covariance matrices, Y is the left hand-side variable of (15) or (16) (i.e., $vech(\varepsilon_t \varepsilon_t')$ or ε_t^2) and W the right hand-side variables (i.e. W_t).

For i.i.d. normally distributed random variables, the likelihood ratio test statistic for the null hypothesis that there exist at least s linear combinations that annihilate $(n-s)$ features in common to these random variables is given by

$$LR_s = -T \sum_{i=1}^s \ln(1 - \hat{\lambda}_i) \quad s = 1, \dots, n, \quad (18)$$

where $\hat{\lambda}_i$ is the i -th smallest eigenvalue of the estimated matrix (17). The test statistic (18) follows asymptotically a $\chi_{(v)}^2$ distribution under the null where $v = snq - s(n-s)$ or $v = sNq - s(N-s)$. The MLE of the matrix δ is given by $\hat{\delta} = [\hat{e}_1, \dots, \hat{e}_s]$, where \hat{e}_i is the eigenvector associated with $\hat{\lambda}_i$ for $i = 1, 2, \dots, s$.

The main drawbacks in using the likelihood ratio test presented above are the normality and independence assumptions on which the test is based on. Indeed, there is huge empirical evidence against normality of squared returns and cross-products. Moreover, when we marginalize a multivariate, strong GARCH process to obtain the process for a subset of variables, a weak GARCH process results, that is *iid*-ness is lost. The likelihood ratio statistic (18) then becomes a quasi-likelihood procedure. Point estimates of the parameters are (strongly) consistent (see e.g. Jeantheau, 1998) and asymptotically normally distributed (see Comte and Lieberman, 2003), but the inverse of the information matrix is not a consistent estimator of the standard errors of the QMLE. Rather one has to use the so-called sandwich formula which pre- and postmultiplies the covariance matrix of the score of the likelihood function by the inverse of the Hessian matrix to consistently estimate the standard-errors of the QMLE. Expressions for this estimator have been derived in the literature (see e.g. Francq and Zakoian,

2007) but they cannot always be easily implemented in practice. It might be sensible to investigate whether bootstrapping standard-errors or test statistics is a sensible procedure in empirical work.

Without a correct implementation of the QMLE estimator, applying the canonical correlation analysis to squared returns and cross-products should lead to an overestimation of the number of common factors. To overcome this problem, Engle and Marcucci (2006) propose to apply the likelihood ratio test on the log of Y and W .⁴ In the case of (16), this would mean testing the rank of \tilde{B} in

$$\ln(\varepsilon_t^2 + \xi) = \gamma_0 + \tilde{B} \ln(W_t + \xi) + v_t, \quad (19)$$

where ξ is a tiny positive constant. Note here that the intuition behind the logarithmic transformation is clear for univariate series and the $\text{diag}(\varepsilon_t \varepsilon_t')$. Indeed we see for instance with $\ln(\varepsilon_{1t}^2) = 2 \ln(u_{1t}) + \ln h_{11t}$ that we retrieve iid-ness properties. Things are much complicated with matrices because the log of the product of two matrices is generally not equal to the sum of the logs of these matrices. We nonetheless evaluate this issue using a naive extension of (19) to the general case (14).⁵

We investigate in this section the small sample behavior of quasi- LR_s test statistics when both the pure variance model ($s = 1, \dots, n$) and the factor model ($s = 1, \dots, n(n+1)/2$) are estimated. The DGP is the BEKK(0, 1, 1) with a single factor, hence common transmission dynamics for the variances but also the covariances. In this setting the pure variance omits to consider these covariances in the estimation of δ and no correction is applied to the LR statistics to account for MA terms.⁶ We take $\Gamma_1 = \sqrt{0.7} \xi (\xi' \xi)^{-1/2}$ where ξ is $(N \times k)$ matrix of independent standard normal random variables (different for each replication) such that $\text{rank}(\Gamma_1) = k$; $\Gamma_0 = \text{chol}(\Omega)$, $\Omega \sim N(0, I_n)$.

We compare the log and the non-log transformation of the squared returns and the squared returns plus the covariances. The entries PV (for pure variance) denote tests on squared returns only; rows with FM (for factor model) denote a system with both squared returns and cross-product moments. We take $T = 500, 1000$ and 2000 . The column labeled $N - 1$ refers to the empirical size; the column labeled N the size unadjusted power. The tables report the rejection frequencies for 1000 replications. The correct number of instruments, i.e. the number of lags l of the dependent variables, is taken in the estimated model with then $l = q = 1$.

It clearly emerges that the best strategy is to take the pure variance model with the logs of the squared returns. Indeed this strategy gives an empirical size close to the 5% nominal one in large sample even when

⁴Engle and Marcucci (2006) mention alternative tests but the majority of them is based on elliptical distributions with possible non-zero excess kurtosis and finite fourth order moments. See Gunderson and Muirhead (1997) and Yuan and Bentler (2000) for more details.

⁵If S is a $(N \times N)$ PSD matrix, by the spectral decomposition theorem we have that $S = E \Lambda E'$, where the columns of the $(N \times N)$ orthonormal matrix E correspond to the eigenvectors of S and Λ is a $(N \times N)$ diagonal matrix whose diagonal elements are the eigenvalues of S . Then the matrix logarithm of S , denoted $\text{logm}(S)$, is defined by $\text{logm}(S) = E \ln(\Lambda) E'$. Recall that the logarithm of a diagonal matrix is a diagonal matrix whose diagonal elements are taken in logs.

⁶An adjustment of the eigenvalues for the presence of a MA component (see Tiao and Tsay, 1989) did not give better results. These outcomes have therefore not been included. Moreover, a strategy based on the no ARCH null hypothesis associated with an estimation of the eigenvectors based on partial least squares (see Cubadda and Hecq, 2010 for the development of this approach for the null of no-autocorrelation) underperforms compared to the canonical correlation test statistics. Results can be obtained upon request.

n increases. Not taking logs in the pure variance model leads to very poor results in terms of size and power. For the system with both squared returns and cross-products (i.e. rows FM), there are size distortions with the non-log transformation. Taking logs seems to improve results but this is only a finite n artefact. Increasing the number of series leads to strong size distortions.

Let us then analyze the behavior of the pure variance estimated models on the log of the squared returns when the DGP is a diagonal VEC(0,1). Indeed we have seen that both the factor model and the diagonal model provide quite parsimonious marginal GARCH orders. We consequently investigate now whether some of the proposed strategies allow us to discriminate between these two models. We report in Table 2 the outcomes for different sample sizes for $n = 2, 5, 15$ and $l = 1, 3, 5$ in the estimated models. It emerges that except when T is small and l is large, the reduced rank procedure rejects quite well the presence of any common factors.

5 An illustrative example

The data set was obtained from TickData and consists of daily closing transaction prices for thirty large capitalization stocks from the NYSE, AMEX NASDAQ, covering the period from January 1, 1999 to December 31, 2008 (2489 trading days). The appendix provides a list of ticker symbols and company names.

For the conditional mean we have estimated AR(2) models with daily dummies to capture Monday and Friday effects. For the conditional variance we have run four different specifications. These are the GARCH(1,1), the GARCH(1,2) and two long memory models, namely FIGARCH(1, d , 0) and FIGARCH(1, d , 1). Out of the 30 series, the GARCH(1,1) model is favoured in six cases using both formal likelihood ratio tests and the Hannan-Quinn information criterion. The six returns with no indication of long memory are (and using their acronyms, see Appendix) ABT, BMY, GE, SLB, XOM, XRX.

Therefore we consider these six return series and we apply our proposed test. The implications of FIGARCH type processes is out of the scope of this paper. Consequently the study of the other 24 series is left for further investigations. Table 3 shows the results for the pure variance model on the six squared returns only as well as the system that also includes their fifteen cross-products, hence with 21 series. We have considered for the moment 2 lags. The number $n - k$ or $N - k$ of linear combinations that do not have the GARCH feature gives the number k of factors generating the pure variance or the factor GARCH model. For the pure variance model in logs, a single common ARCH component generating the six returns is found, that is we take $k = 1$. The results would be different for the model in levels and/or when one also stacks the covariances. However, the simulation results of Section 4 revealed that the pure variance model for the logs of squared returns yields better results.

6 Conclusion and further research

This paper studies the orders of the marginal weak GARCH processes implied by a multivariate GARCH model. We see that except for some coincidental situation the marginal models are non-parsimonious. We then show

Table 1: Size and power of common volatility for the non-log version tests statistics

			non logs- LR		logs- LR	
			$N - 1$	N	$N - 1$	N
$n = 2$	$T = 500$	PV	14.5	99.9	4.6	99.2
		FM	19.4	100	5.7	100
	$T = 1000$	PV	17.8	100	6.3	100
		FM	25.3	100	5.6	100
	$T = 2000$	PV	20.9	100	5.3	100
		FM	29.5	100	5.6	100
$n = 5$	$T = 500$	PV	30.9	99.9	3.4	80.9
		FM	52.7	98.8	11.2	97.8
	$T = 1000$	PV	38.0	100	4.0	99.3
		FM	63.0	100	10.3	100
	$T = 2000$	PV	48.4	100	6.8	100
		FM	73.5	100	9.0	100
$n = 15$	$T = 500$	PV	41.6	91.3	3.5	37.9
		FM	100	100	100	100
	$T = 1000$	PV	45.5	99.4	3.9	59.5
		FM	100	100	100	100
	$T = 2000$	PV	57.0	100	4.3	88.2
		FM	100	100	100	100

Note: The DGP has 1 source of volatility. Hence, the columns $N - 1$ report the empirical size while the column N report the (size unadjusted) power. PV refers to the pure variance estimated model while FM refers to the factor model in which we also stack the covariances. $l = 1$ in the estimated model. Non-logs LR columns refer to the likelihood ratio test on the vech or the diag of $\varepsilon_t \varepsilon_t'$; logs-LR columns refer to the likelihood ratio test on the vech or the diag of the matrix logs of $\varepsilon_t \varepsilon_t'$. Nominal size is 5%.

Table 2: Power of the PV common volatility for the log version tests statistics for the diagonal BEKK

		$l = 1$			$l = 3$			$l = 5$		
		$T=500$	1000	2000	$T=500$	1000	2000	$T=500$	1000	2000
$n = 2$	$s = 1$	99.8	100	100	98.8	100	100	96.3	100	100
	$s = 2$	100	100	100	100	100	100	100	100	100
$n = 5$	$s = 1$	95.1	100	100	56.8	99.4	100	34.2	97.1	100
	$s = 2$	99.6	100	100	93.3	100	100	83.2	100	100
	$s = 3$	100	100	100	99.7	100	100	98.1	100	100
	$s = 4$	100	100	100	100	100	100	99.9	100	100
	$s = 5$	100	100	100	100	100	100	100	100	100
$n = 15$	$s = 1$	5.0	94.2	100	0.0	3.3	98.5	0.0	0.2	86.1
	$s = 2$	11.9	99.7	100	0.0	24.1	100	0.0	3.8	99.4
	$s = 3$	27.6	100	100	0.0	58.3	100	0.0	19.7	100
	$s = 4$	49.7	100	100	0.1	87.9	100	0.0	51.7	100
	$s = 5$	73.4	100	100	2.4	98.0	100	0.7	81.9	100
	$s = 6$	89.7	100	100	14.5	97.7	100	4.1	96.4	100
	$s = 7$	96.3	100	100	39.0	100	100	22.0	99.4	100
	$s = 8$	99.2	100	100	71.1	100	100	53.6	100	100
	$s = 9$	100	100	100	91.4	100	100	81.1	100	100
	$s = 10$	100	100	100	98.7	100	100	96.7	100	100
	$s = 11$	100	100	100	100	100	100	99.7	100	100
	$s = 12$	100	100	100	100	100	100	100	100	100
	$s = 13$	100	100	100	100	100	100	100	100	100
	$s = 14$	100	100	100	100	100	100	100	100	100
	$s = 15$	100	100	100	100	100	100	100	100	100

Note: The DGP has n source of volatility within n variables. This is the diagonal VEC(0,1). We report the power of the pure variance estimated model with the log version.

Table 3: P-Value of LR tests (2 lags) for both the pure variance and the factor model

H_0	Pure Variance		H_0	Factor Model	
	non log	logs		non logs	logs
$k \leq 6$	0.936	0.983	$k \leq 21$	0.504	1.000
$k \leq 5$	0.051	0.983	$k \leq 20$	0.773	1.000
$k \leq 4$	0.000	0.984	$k \leq 19$	0.812	1.000
$k \leq 3$	0.002	0.883	$k \leq 18$	0.284	1.000
$k \leq 2$	0.503	0.443	$k \leq 17$	0.064	1.000
$k \leq 1$	0.000	0.000	$k \leq 16$	0.358	1.000
			$k \leq 15$	0.240	1.000
			$k \leq 14$	0.111	1.000
			$k \leq 13$	0.159	1.000
			$k \leq 12$	0.350	1.000
			$k \leq 11$	0.148	1.000
			$k \leq 10$	0.002	1.000
			$k \leq 9$	0.004	1.000
			$k \leq 8$	0.001	1.000
			$k \leq 7$	0.007	0.999
			$k \leq 6$	0.000	0.998
			$k \leq 5$	0.002	0.921
			$k \leq 4$	0.011	0.414
			$k \leq 3$	0.126	0.015
			$k \leq 2$	0.046	0.000
			$k \leq 1$	0.000	0.000

that multivariate models with co-movements in the conditional volatility and cross products explain why we can obtain GARCH with small orders in empirical work. This result would plead for looking at individual series prior to a multivariate modelling.

We propose a simple strategy to detect the presence of such GARCH co-movements and we apply it to daily returns from the NASDAQ. It emerges that when working with the six returns (out of 30) having parsimonious univariate GARCH specifications we detect the presence of a single factor generating the volatility.

From our Monte Carlo study, the best strategy consists in applying the method proposed by Engle and Marcucci (2006) for the logs of the squared returns (i.e. the pure variance model), omitting therefore the cross-products (although they might matter in theory).

Many issues are currently under investigation. First we have focussed on a limited number of multivariate models. We have not covered stochastic volatility models, or alternative multivariate models (GO-GARCH, DECO,...). Second, we have still to study the properties of the tests proposed using appropriate estimators of the asymptotic standard-errors of the QMLE and investigate the behavior of bootstrap procedures to estimate standard-errors and to test for the number of common factors in multivariate weak GARCH processes.

Third, our test statistics seems to be tailored for multivariate GARCH(0, q) but we have not extensively investigated in our Monte Carlo study its use in more general multivariate GARCH(p, q) although results for the marginalization orders of the multivariate GARCH(p, q) are given in this paper. Whether this multivariate modelling is more appropriate than a GARCH(0, q) is an empirical issue. In order to work in this framework, and for separating the MA from the AR part possibly with matrices with different left null spaces, we believe that another testing strategy should be implemented. A promising procedure we leave for further investigations relies on realized variance and covariances.

7 References

- BABA, Y., R.F. ENGLE, D.F. KRAFT, AND K.F. KRONER (1989), Multivariate Simultaneous Generalized ARCH, DP 89-57 San Diego.
- BAUWENS, L., S. LAURENT AND J.V.K. ROMBOUTS (2006), Multivariate GARCH Models: a Survey, *Journal of Applied Econometrics* 21, 79–109.
- BILLIO M., M. CAPORIN AND M. GOBBO (2003), Block Dynamic Conditional Correlation Multivariate GARCH Models, Working paper 0303 GRETA, Venice.
- COMTE, F. AND O. LIEBERMAN (2003), Asymptotic Theory for Multivariate GARCH Processes, *Journal of Multivariate Analysis* 84, 61-84.
- CUBADDA, G., HECQ A. AND F.C. PALM .(2008), Macro-panels and Reality, *Economics Letters*, 99, 537-540.
- CUBADDA, G., HECQ A. AND F.C. PALM (2009), Studying Interactions without Multivariate Modeling, *Journal of Econometrics*, 148, 25-35.
- CUBADDA, G. AND A. HECQ (2010), Testing for Common Autocorrelation in Data Rich Environments,

forthcoming in *Journal of Forecasting*.

DROST, F.C. AND T.E. NIJMAN (1993), Temporal Aggregation of GARCH Processes, *Econometrica* 61, 909-927.

ENGLE, R. F., V. NG AND M. ROTHCHILD (1990), Asset Pricing With a Factor ARCH Covariance Structure: Empirical Estimates of Treasury Bills, *Journal of Econometrics*, 45, 213-238.

ENGLE, R. F. AND B. KELLY (2008), Dynamic Equicorrelation, Stern School of Business, WP NYU.

ENGLE, R. F. AND S. KOZICKI (1993), Testing for Common Features (with Comments), *Journal of Business and Economic Statistics*, 11, 369-395.

ENGLE, R. F. AND J. MARCUCCI (2006), A Long-run Pure Variance Common Features Model for the Common Volatilities of the Dow Jones, *Journal of Econometrics*, 132, 7-42.

FRANCQ, C. AND J.-M. ZAKOIAN (2007), HAC Estimation and Strong Linearity Testing in Weak ARMA Models, *Journal of Multivariate Analysis* 98, 114-144.

JEANTHEAU, T. (1998), Strong Consistency of Estimators for Multivariate ARCH Models, *Econometric Theory* 15, 70-86.

NIJMAN, T. AND E. SENTANA (1996), Marginalization and Contemporaneous Aggregation in Multivariate GARCH Processes, *Journal of Econometrics*, 71, 71-87.

SILVENNOINEN, A. AND T. TERASVIRTA (2009), *Multivariate GARCH Models*, Handbook of Financial Time Series, Ed. Andersen, T.G. and Davis, R.A. and Kreiss, J.P. and T. Mikosch, Springer.

TIAO, G. C. AND R. S. TSAY (1989), Model Specification in Multivariate Time Series (with Comments), *Journal of Royal Statistical Society, Series B*, 51, 157-213.

ZELLNER, A. AND F.C. PALM (1974), Time Series Analysis and Simultaneous Equation Econometric Models, *Journal of Econometrics* 2, 17-54.

ZELLNER, A. AND F.C. PALM (1975), Time Series and Structural Analysis of Monetary Models of the US Economy, *Sanhya: The Indian Journal of Statistics*, Series C 37, 12-56.

ZELLNER, A. AND F.C. PALM (2004), *The Structural Econometric Time Series Analysis Approach*, Cambridge University Press.

8 Stocks used in the empirical application

Symbol	Issue name	Symbol	Issue name
AAPL	APPLE INC	JNJ	JOHNSON &JOHNSON
ABT	ABBOTT LABORATORIES	JPM	JP MORGAN CHASE
AXP	AMERICAN EXPRESS CO	KO	COCA COLA CO
BA	BOEING CO	LLY	ELI LILLY & CO
BAC	BANK OF AMERICA	MCD	MCDONALDS CORP
BMJ	BRISTOL MYERS SQ	MMM	3M COMPANY
BP	BP plc	MOT	MOTOROLA
C	CITIGROUP	MRK	MERCK & CO
CAT	CATERPILLAR	MS	MORGAN STANLEY
CL	COLGATE-PALMOLIVE CO	MSFT	MICROSOFT CP
CSCO	CISCO SYSTEMS	ORCL	ORACLE CORP
CVX	CHEVRON CORP	PEP	PEPSICO INC
DELL	DELL INC	PFE	PFIZER INC
DIS	WALT DISNEY CO	PG	PROCTER & GAMBLE
EK	EASTMAN KODAK	QCOM	QUALCOMM
EXC	EXELON CORP	SLB	SCHLUMBERGER N.V.
F	FORD MOTOR CO	T	AT&T CORP
FDX	FEDEX CORP	TWX	TIME WARNER
GE	GENERAL ELEC	UN	UNILEVER N V
GM	GENERAL MOTORS	VZ	VERIZON COMMS
HD	HOME DEPOT INC	WFC	WELLS FARGO & CO
HNZ	H J HEINZ CO	WMT	WAL-MART STORES
HON	HONEYWELL INTL	WYE	WEYERHAEUSER CO
IBM	INTL BUS MACHINE	XOM	EXXON MOBIL
INTC	INTEL CORP	XRX	XEROX CORP